Navigation Support for Students who are Blind in Accessing Discussion Boards

Fabrizio Antonelli
Jong Wook Kim
K. Selçuk Candan
Maria Luisa Sapino

Universita’ di Torino, Italy
Arizona State University, USA

Consider an educational web site
Educational web site

- Complex Navigation Tools
- Frame heavy design
- Hierarchically organized content with high navigation req.

Motivation

- Very good if you (the student) have sight....

- ......pretty bad, if you don’t.

Presented by Maria Luisa Sapino
iCare-Assistant

- Dynamic adaptation and guidance
- Transpareently reduce navigation load!!!

Adaptation is performed transparently at a higher level

Existing content are not modified

A specific example: discussion board

<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>buzz proj.</td>
<td>Vander, Ryan</td>
<td>Tue May 26, 2004 9:21 am</td>
</tr>
<tr>
<td>Re: buzz proj.</td>
<td>True, Thomas</td>
<td>Thu May 27, 2004 7:53 pm</td>
</tr>
<tr>
<td>Re: buzz proj.</td>
<td>Vander, Ryan</td>
<td>Sat May 29, 2004 2:08 pm</td>
</tr>
<tr>
<td>Re: buzz proj.</td>
<td>Grain, Robert</td>
<td>Sun May 30, 2004 6:10 pm</td>
</tr>
<tr>
<td>Assignment 4</td>
<td>Rodriguez, Luisa</td>
<td>Thu May 27, 2004 3:04 pm</td>
</tr>
<tr>
<td>Report for Assig. 4</td>
<td>True, Thomas</td>
<td>Thu May 27, 2004 7:57 pm</td>
</tr>
<tr>
<td>Re: Report for Assig. 4</td>
<td>Candan, Kasim</td>
<td>Mon May 31, 2004 12:07 am</td>
</tr>
<tr>
<td>Assignment #4</td>
<td>Atilla, John</td>
<td>Fri May 28, 2004 10:41 pm</td>
</tr>
<tr>
<td>Re: Assignment #4</td>
<td>Candan, Kasim</td>
<td>Mon May 31, 2004 12:19 am</td>
</tr>
<tr>
<td>Questions on #4</td>
<td>Roosevelt, Daniel</td>
<td>Sat May 29, 2004 11:00 pm</td>
</tr>
<tr>
<td>Re: Questions on #4</td>
<td>Candan, Kasim</td>
<td>Mon May 31, 2004 12:23 am</td>
</tr>
<tr>
<td>Re: Questions on #4</td>
<td>Ray, Luisa</td>
<td>Mon May 31, 2004 10:54 pm</td>
</tr>
<tr>
<td>Re: Questions on #4</td>
<td>Home, Chris</td>
<td>Tue Jun 1, 2004 12:23 am</td>
</tr>
<tr>
<td>Report Length</td>
<td>True, Thomas</td>
<td>Tue Jun 1, 2004 11:39 am</td>
</tr>
<tr>
<td>Assignment #4</td>
<td>Bird, Sarah</td>
<td>Tue Jun 1, 2004 9:14 pm</td>
</tr>
</tbody>
</table>
iCare-Assistant

- Dynamic adaptation and guidance (context and task)
- Transparently reduce navigation load!!!

Simple Book View
Number of Alternatives: 1
Number of Steps to Target: ???

Naive Content-based View
Number of Alternatives: 3
Number of Steps to Target: 2

Assisted Context-based View
Number of Alternatives: 1
Number of Steps to Target: 1

How can we provide adaptive access to discussion board messages?

buzz proj. Vander, Ryan Tue May 26, 2004 9:21 am
Re: buzz proj. True, Thomas Thu May 27, 2004 7:05 pm
Re: buzz proj. Vander, Ryan Sat May 29, 2004 2:08 pm
Assignment 4 Rodriguez, Luisa Thu May 27, 2004 3:04 pm
Report for Assig. 4 True, Thomas Thu May 27, 2004 7:57 pm
Re: Report for Assig. 4 Candan, Kasim Mon May 31, 2004 12:07 am
Assignment #4 Atilla, John Fri May 28, 2004 10:41 pm
Re: Assignment #4 Candan, Kasim Mon May 31, 2004 12:19 am
Questions on #4 Roosevelt, Daniel Sat May 29, 2004 11:00 pm
Re: Questions on #4 Candan, Kasim Mon May 31, 2004 12:23 am
Re: Questions on #4 Ray, Luisa Mon May 31, 2004 10:34 pm
Re: Questions on #4 Home, Chris Tue Jun 1, 2004 12:23 am
Report Length True, Thomas Tue Jun 1, 2004 11:39 am
Assignment # 4 Bird, Sarah Tue Jun 1, 2004 9:14 pm

Presented by Maria Luisa Sapino
Can we properly interpret messages???

Messages are too short for indexing

Leverage context
Our solution: Navigational adaptation and annotation

Chicken-and-Egg problem

- Rich keywords are needed for understanding topic boundaries
- Without boundaries, we can’t identify which keywords should belong to which messages
Chicken-and-Egg problem

- Rich keywords are needed for understanding topic boundaries
- Without boundaries, we can’t identify which keywords should belong to which message
- Our approach: first try to understand how the discussion evolve...don’t worry about the topics themselves

How do contexts/topics evolve???

Presented by Maria Luisa Sapino
Problem Statement

- Find entry points to *new*, *general*, or *specific* topics within a single hierarchy of dynamically evolving web content
  - algorithms for identifying how the topic content of a discussion board evolves
  - identifying coherent segments of the content

- …use this for indexing and guidance.

Segmenting a Single Message Chain

Presented by Maria Luisa Sapino
Two step segmentation of a message

- **Step 1**: new topic or not??
  (Low-granularity)
  - Identify whether the message is an unrelated topic or not

- **Step 2**: if not new, then what??
  (High-granularity)
  - Determine whether the message is more specific or more general than the previous messages

---

**Step 1: Identifying New Topic Boundaries**

- **Challenge**
  - consecutive messages of the same topic may be of different length, style, and content
  - original posting can be included in replies as quotations (out of context)
Quotations

Quote: two different nodes have the same ad-labels? I found in the Quote data file produced by JK's code, there exists two different Quote nodes with the same ad-labels and different ses-labels!

..this is curious... can you please give us more details regarding this case?

A quotation-reply pair can provide "context" pointers across messages.

Step 1: Identifying New Topic Boundaries

- the aggregate frequency of the keyword

\[ freq_i = freq_{i,0} + freq_{i,d} \]

where
- \( freq_{i,0} \) is the frequency of the keyword \( k_i \) in the message excluding the quotations
- \( freq_{i,d} \) is its frequency in \( d \)-level quotations
- \( imp(d) \) is the impact factor of the quotations that are of depth \( d \)

Keywords in quotations

Presented by Maria Luisa Sapino
Step 1: Identifying New Topic Boundaries

- Once a keyword weight vector is computed for a message,
  - the cosine similarity,
  - Hellinger distance, or
  - Kullback-Leibler divergence
  can be used to classify the input message as new topic or same topic

Step 2: Segmentation based on Specialization / Generalization

- **Specialization**
  - focus on a specific aspect of the common theme
  - ....more constraints on the topic (“and”)

- **Generalization**
  - take the discussion to a more general platform
  - ...less constraints on the topic (“or”)

Presented by Maria Luisa Sapino
Generalization/Specialization

In the keyword space \(<k_1,k_2>\), assume
- \(m_A = \ldots k_1 \ldots k_2 \ldots k_2 \ldots k_1 \ldots\)
- \(m_B = \ldots k_1 \ldots k_1 \ldots k_1 \ldots k_1 \ldots\)
- \(m_O = \ldots k_3 \ldots k_8 \ldots k_6 \ldots k_9 \ldots\)

Message representation:
- \(m_A = < n_1, n_2>\)
- \(m_B = < n_1, 0>\)
- \(m_O = < 0, 0>\)

Step 2: Segmentation based on Specialization / Generalization
**Specialization / Generalization**

In the keyword space \(<k_1, k_2>\)

\[ m_0 = "\ldots k_5 \ldots \ldots k_7 \ldots \ldots k_4 \ldots k_9 \ldots" \]

\[ m_0 = <0, 0> \equiv (\neg k_1 \land \neg k_2) \equiv \neg (k_1 \lor k_2) \]

\( m_A \) more general than \( m_B \) \( \rightarrow \)

\[ \Delta_{mA} = \text{dist}(m_A, O) > \Delta_{mB} = \text{dist}(m_B, O) \]

---

**Step 2: Segmentation based on Specialization / Generalization**

- the degree of *generalization*,

\[ G_{AB} = \frac{\Delta_{mA}}{\Delta_{mB}} \]

---

Presented by Maria Luisa Sapino
Step 2: Segmentation based on Specialization / Generalization

- the degree of *specialization*,
  \[ S_{AB} = \frac{\Delta m_B}{\Delta m_A} \]

Can we leverage quotation
- anchored parts
  - quotation messages from the parent and ancestor + the parts of the message to reply to these quotations
- free parts
  - part which is not immediately associated with the parent of ancestor quotations
Step 2: Segmentation based on Specialization / Generalization

- the degree of generalization,

\[
\frac{1}{N} \begin{array}{c}
\text{free part} \\
\text{anchored part}
\end{array}
\]

where
- \(N\) is the number of keywords in the message
- \(d_{\text{free}}\) is the free part and \(n_{\text{free}}\) is # of keywords in this part
- \(q_i, d_i\) is the \(i^{th}\) anchored quotation-reply pair and \(n_{\text{anch},i}\) is # of keywords in this pair
- \(D_{\text{par}}\) is the parent message
Segmenting a Hierarchy of Messages

- Independent message
  - two separate replies to a single message are independently created from each other

- The two-step segmentation process is repeated, following each chain of the hierarchy

Context-sensitive message classification

- Goal: to extract knowledge about the structural inter-dependencies among messages
  - For navigation purposes
  - For direct use in queries

Presented by Maria Luisa Sapino
Message classification

- Question
- Direct Answer
- Indirect Answer
- Event Announcement
- Recommendation
- Other

Fuzzy classification scheme!

\[ m \rightarrow \{ <Q, \text{score}_Q>, <DA, \text{score}_{DA}>, <IA, \text{score}_{IA}>, <EN, \text{score}_{EN}>, <R, \text{score}_R>, <O, \text{score}_O> \} \]
Classification module

- Step 1: content based classification

Rule example

(defrule there_is_link
  (declare (salience 80))
  (?f1 <- (score (post_ID ?id) (type indirect_answer) (value ?v1)))
  (?f2 <- (score (post_ID ?id) (type recommendation) (value ?v2)))
  (url_link (post_ID ?id))
  (not (done (op there_is_link) (arg1 ?id)))
  (increment (type indirect_answer) (value ?i1))
  (increment (type recommendation) (value ?i2))
  =>
  (modify ?f2 (value (+ ?v2 (* 3 ?i2))))
  (modify ?f1 (value (+ ?v1 (* 3 ?i1))))
  (assert (done (op there_is_link) (arg1 ?id))))
### Contribution Matrix

<table>
<thead>
<tr>
<th>CONDITION (RULE)</th>
<th>QUESTION (Q)</th>
<th>DIR_ANS. (DA)</th>
<th>IND_ANS. (IA)</th>
<th>RECOMMEN. (R)</th>
<th>ANNOUN.(N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of QUESTION MARKS &gt; 0</td>
<td>□</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time reference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>□</td>
</tr>
<tr>
<td>Similarity between the hyperlink keywords and the...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The symbols “■” and “□” indicate a score to be modified when the rule is fired.
- “■” indicates that the increment value for the score is an absolute real value
- “□” indicates that the increment value for the score is variable (for instance depends on the similarity value)

### Classification module

- **Context Based Classification**

  - m classified as Question (without a corresponding answer) → similarity based search for OTHER messages, among the descendants, that can be recognized as answer to m
Classification module

- Context Based Classification

  - m classified as Answer (IA or DA) (without a corresponding question) → similarity based search for OTHER messages, among the ancestors, that can be recognized as answer to m. The score of m as an answer is possibly updated (since it is supported by the presence of the question)
Experiments: Segmentation

- Experiment Setup
  - 20 discussion threads
  - a total of 368 messages
  - average thread depth of 12.45
  - average quotation depth of 1.3
    - 86% of quotations are from the parent
  - 5 user study
Experiments

- Evaluation Criteria

\[ \text{success \_ rate} = \frac{\sum_{m=\text{messages}} 1 - \text{error \_ weight}(m)}{\text{number \_ of \_ messages}} \times 100 \]

Presented by Maria Luisa Sapino
Experiments

- Success rate

<table>
<thead>
<tr>
<th>Success rate (%)</th>
<th>Undiff.</th>
<th>Low-only</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>79.06%</td>
<td>87.31%</td>
<td>83.19%</td>
</tr>
</tbody>
</table>

Experiments

- Effect of quotations

<table>
<thead>
<tr>
<th>Quot. weights</th>
<th>Success rate (%)</th>
<th>Undiff.</th>
<th>Low-only</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off</td>
<td></td>
<td>72.57%</td>
<td>85.25%</td>
<td>78.00%</td>
</tr>
<tr>
<td>On</td>
<td></td>
<td>79.06%</td>
<td>87.31%</td>
<td>83.19%</td>
</tr>
</tbody>
</table>
Experiments: Annotation

- 61 messages
- 8 assessors

- User classification
  Q: 31%  IA: 8%  DA: 20%  EN: 13%  R: 13%  O:15%

- System classification
  Q: 34%  IA: 10%  DA: 18%  EN: 11%  R: 13%  O:13%

User vs. System classification

<table>
<thead>
<tr>
<th>User / System</th>
<th>Q</th>
<th>DA</th>
<th>IA</th>
<th>EN</th>
<th>R</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DA</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IA</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>EN</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>O</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

Individual precision:
Q: 1.00, IA: 0.80, DA: 0.83, EN: 0.88, R: 0.88, and O: 0.76.
User vs. System classification

<table>
<thead>
<tr>
<th>User / System</th>
<th>Q</th>
<th>DA</th>
<th>IA</th>
<th>EN</th>
<th>R</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DA</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IA</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>EN</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>O</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

Overall precision: 89%

Conclusions

- We presented algorithms
  - for identifying how the hierarchical content of an educational discussion board grows through generalizations and specializations; and
  - for annotating navigation pathways.

- The algorithms are deployed in iCare-Assistant for blind students, but are applicable for developing other educational tools.
Related Work

- Topic Segmentation (text)
- Topic Distillation (web)
- Topic Tracking (news)
- Video Segmentation (video sequences)

- Adaptive Hypermedia
- Adaptive and Assistive Technologies